Volume 15 Number 6 1987 Nucleic Acids Research

Charons 36 to 40: multi enzyme, high capacity, recombination deficient replacement vectors with polylinkers and polystuffers

Ian S.Dunn[1] and Frederick R.Blattner*

Laboratory of Genetics, University of Wisconsin, Madison, WI 53706, USA

Received December 18, 1986; Revised and Accepted February 11, 1987

ABSTRACT

New phage λ based cloning vectors, Charons 36-40, have been constructed which allow cloning of large (up to 24 kb) DNA fragments with up to sixteen cloning enzymes. Several of these could not be used previously with λ vectors. Clones produced with these vectors can be propagated under recombination deficient conditions. A novel polystuffer method has been developed that permits vector arms to be purified by simple precipitation and which allows reliable identification of clones that have reincorporated any part of the stuffer. Three of the vectors are available with amber mutations in essential genes.

INTRODUCTION

Engineering of the genome of bacteriophage λ has led to development of DNA cloning vectors of steadily increasing utility (for general reviews of λ biology and vectors see 1,2). Improved λ vectors have been described recently (3,4,5). In this communication we describe further developments on this line resulting in a series of improved cloning vectors, Charons 36-40. The relevant cloning features of these phage are:

(1) Cloning experiments can now be performed with up to 16 different restriction enzyme specificities. Four new mutations were isolated that remove restriction sites from the essential portions of λ and an extended polylinker was created for these and other cloning sites that are naturally absent from λ. (A seventeenth enzyme is available but its use obviates the ability of the vectors to grow on *recA*$^-$ strains.)

(2) The polylinkers at the two cloning sites are in inverse orientation, allowing cloned fragments to be excised with any enzymes exterior to the one chosen for insertion, including many that could not be used for the initial cloning because of sites in the essential portion of the vector.

(3) BamH I is at the interior end of the polylinker so that libraries can be constructed with Mbo I or Sau 3A cut target DNA and ligated to the BamH I cut vector DNA. The resulting cloned fragments can be excised by any of the other restriction enzyme sites in the polylinker, even if the BamH I site is eliminated by the cloning process.

(4) Capacity was increased by development of somewhat shorter right and left arms than were available in previous vectors owing to deletion of nonessential genes beta, a portion of exo and a portion of lom. This results in the highest capacity yet achieved in a general purpose lambda replacement vector (up to 24 kbp. for Ch 38-40).

(5) Removal of stuffer fragments has been made extremely simple and efficient in two of the new vectors by the use of a polystuffer. This consists of many tandem copies of a short DNA segment so that cutting with a single enzyme reduces the entire stuffer to small pieces that can be removed by polyethylene glycol (PEG) precipitation. The lac operator is present on each copy which allows reliable identification of clones that by accident include any of the stuffer segments.

(6) The cloning vectors are particularly suited for cloning DNA segments that are unstable with respect to general recombination, for example those containing direct or indirect repeats. The phages carry the lambda gam gene on the right arm. Gam is an inhibitor of the rec BC nuclease so its expression results in a recombination deficient phenotype when propagation is on wild type *Escherichia coli* and permits propagation on recA *E. coli* for even lower levels of recombination.

(7) All the DNA of the vectors (including polystuffer but not ordinary stuffer fragments) is from sequenced sources. This allows convenient mapping of inserted target DNA, since all vector derived fragments can be readily identified by computer scan.

MATERIALS AND METHODS

Bacterial phage and strains

K802 is *hsdR*$^-$, *hsdM*$^+$, *gal*$^-$, *met*$^-$, *supE*; LE392 is *hsdR*$^-$, *hsdM*514, supE44, supF58, LacZ$^+$, galK2, GalT22, metB1, trpR55; ED8767 is *galK2*, *galT22*, *recA56*, *metB1*, *hsdS3*, λ^-, supE44, supF58, LacZ$^+$. Chs 33, 34 and 35 have been described previously (6). λFAC101 was kindly provided by F. Lawyer and D. Gelfand (Cetus Corp.). It has a Sal I site at position 19159 where wild-type λ has an Alu I site. The way this was accomplished was as follows: Plasmid vector π AN7 (described in New England Biolabs catalog, 33), which carries Sup F, was cut with Sal I, the ends filled in with Pol I and then digested with BamH I. The Sau 3A to Alu I fragments of λ flanking the Alu I site on the left (18783-19159) was isolated and cloned into the plasmid, creating a reconstituted Sal I site where the Alu I site was joined to the filled in Sal I site. Charon 30A (7) was passaged through the plasmid-bearing strain and plated on an Su^0 strain to select for a recombinational insertion of the plasmid and its suppressor tRNA into the phage (generating a new Sal I site at position 19159).

Phage growth

λ phage were propagated in liquid culture by the PDS method (8) as follows: one ml of an overnight culture of a suitable host strain, generally LE392, grown in NZC medium (GIBCO) is adsorbed for 15 minutes at 37 C with 2 x 10^7 phage (multiplicity of infection of 0.01). After the adsorption, 2 ml of NZC broth is added to the cells which are then used to inoculate 1 liter of NZC broth in a 2 liter flask. This is incubated on a rotary shaker at 250 rpm at 37 C for 12-16 hr until visible lysis has occurred (the culture should be regularly monitored after 12 hr to find the earliest time with good lysis, evidenced by coagulated debris and fairly clear medium). One ml of chloroform is added, the flasks shaken 15 min, and lysates are then cleared by centrifugation (5k rpm/30 min/ 4 C in an RC3B centrifuge) and samples taken for titering. Phage are then precipitated with polyethylene glycol 8000 (PEG, 70 g/l) and NaCl (60 g/l) for at least 2 hr at 0 C (ice-water bath), and pelleted by centrifugation at 5k rpm for 30 min at 4 C in the RC3B centrifuge.

The PEG pellet is resuspended in 30 ml of CsCl (ρ = 1.55,in 10 mM $MgCl_2$, 100 mM NaCl and 50 mM Tris pH 7.4;) and banded in an ultracentrifuge (60Ti rotor at 25-30K rpmfor 18-24 hr at 18 C). Phage are further purified by rebanding (CsCl, ρ =1.50, SW50.1 or AH650 swing-out rotor, 25-30K rpm, 18-24 hr, 18 C). Final collected bands are stored in the CsCl.

Oligonucleotides

Oligonucleotides were synthesized by helpful friends with DNA synthesis machines: John Rossi and Bruce Kaplan of City of Hope, Duarte, CA., Richard Malvarca, Cistron, Pine Brook, NJ. and Ron Niece, UW Biotechnology Center, Madison, WI.

Recombinant DNA methods

General DNA manipulations including restriction enzyme digests, ligations for cloning, λ packaging and gel electrophoresis were performed by standard methods as described (11). Chromogenic detection of lac operator used "XG" plates as described in (8). Computation was performed on a DNASTAR laboratory microcomputer system (12).

Hybridizational screening and phage "lifts" were done as described (13), except special hybridization conditions were used for the identification of Ch 36 candidates. The probe was a section of the pUC19 polylinker that was 5'-end labelled at the EcoR I site, recut with Sma I, and purified. This short fragment is perfectly homologous to the relevant region of the Ch 36 polylinker, but has mismatches with the original linker of Ch 35, allowing discrimination between the parental phage and the new vector candidates. Phage plaques were lifted onto nitrocellulose as usual (13). Hybridizations were done in 5 x Denhardt's solution, 6 x SSC, 0.1 % SDS with

2 x 10^5 cpm/ml of 5'-labelled probe at 30 C for 16 hr. Suitable washes (once with 2 x SSC, 0.1 % SDS; twice with 1 x SSC, 0.1 % SDS; all at 30 C for 30 min each) allowed adequate discrimination between the polylinker sequences.

Directionalized polymerization (polystuffer construction)

The basic strategy pursued was to find a restriction fragment having ends produced by two different enzymes with compatible overhangs which when ligated together cannot be cut by either of the original enzymes. Examples are BamH I /Bgl II; Xho I /Sal I; Xba I /Spe I /Nhe I /Avr II. After ligation of such a fragment, a head to tail ligation product can be selected by redigestion with both the restriction enzymes as only the desired polymer will lack the original sites. Ligation was done in PEG with high concentrations of sodium ions (9,10). The enzymes used, BamH I and Bgl II, retain activity under the ligation conditions used. Thus, it was feasible to include the restriction enzymes in the ligation mixture so that both ligation and recutting occurred concurrently. Ligations were done in 30 μ l volumes with 0.5-1.0 μ g of purified fragment, in 50 mM Tris pH7.4, 7 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 100 μ g/ml bovine serum albumin (nuclease-free), 150 mM NaCl, 10% PEG, 10 units of BamH I and Bgl II, and 1.5 Weiss units of T4 DNA ligase (BRL). After 2 hr at 37 C, most of the input DNA was converted to species of >50 kb. Success of the directionalization procedure could be demonstrated by the failure of purified high-molecular weight product to cut with either BamH I or Bgl II. As the polymerized fragments used here lacked internal Xho II sites, the only Xho II sites occur at correct BamH I/Bgl II junctions. This was useful for demonstrating that the head-to-tail polymerization did not involve aberrant ligations. Also, it was easier in practice to control the size distribution of the products by allowing ligation to proceed to completion and then partially redigesting the product with Xho II.

PEG precipitation for removal of phage polystuffers

Phage vector DNA (Ch 39, 39A, 40 or 40A) is digested with the cloning enzyme of choice, and with Nae I to reduce the stuffer to 235 or 540 bp blunt-end pieces. Based on the method of Lis (14), DNA (50-100 ug/ml), in 0.5 M NaCl, 10 mM Tris pH 7.4, 1 mM EDTA is adjusted with PEG 8000 (Sigma) added to 5% (from a 40% stock solution). After incubation at room temperature for 1-2 hr the preparation is centrifuged for 5 min in a microcentrifuge to pellet phage arms. To remove trapped stuffer fragments, the mixture is resuspended in 200 μ l of the same PEG/salt/Tris solution and incubated at room temperature another 30 min and centrifuged as before. The pellet is redissolved in 200 μl 0.2M NaCl and extracted once with phenol, and once with chloroform. Finally, the DNA is precipitated with 2 volumes of ethanol and washed with 70 ethanol. The dried DNA, now free from PEG, may be resuspended in Tris/EDTA.

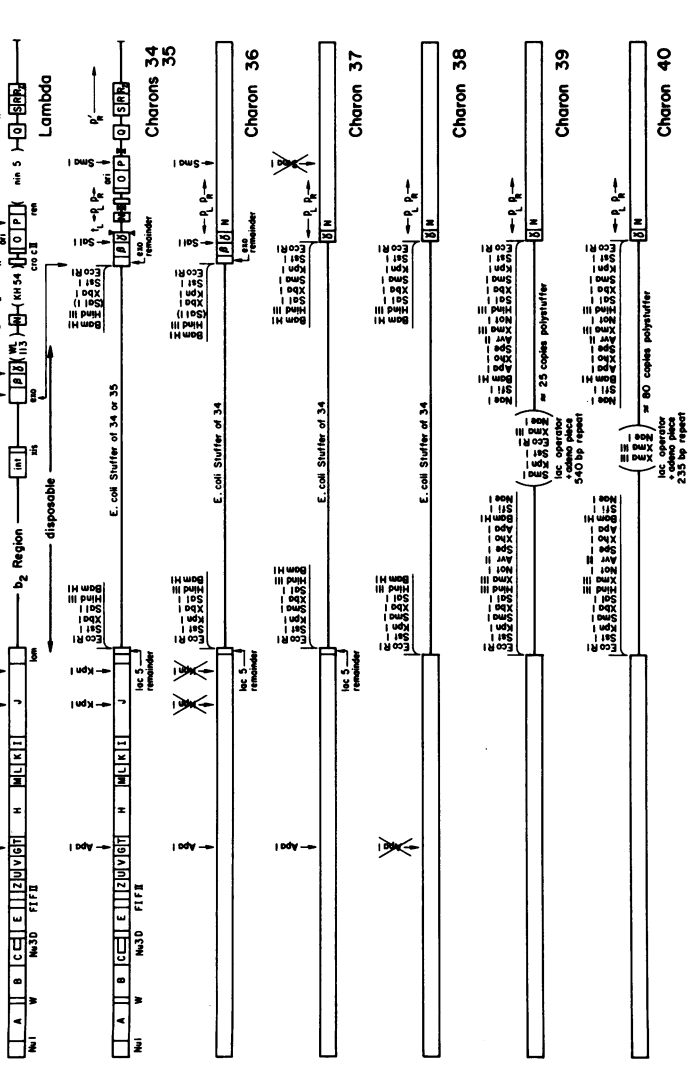

Fig 1: The maps of the new Charon vectors are presented in relation to the wild-type lambda sequence from which they were derived. The disposable region represents DNA of lambda that can be replaced by foreign DNA. Charons 34 and 35 and 39 and 40 are the same except for the stuffers. The three deletions of the right arm, WL113, KH54 and nin5 are indicated with darkened triangles on the Charon 34-35 map. Restriction sites that are crossed off have been removed by mutation. P_L and P_R and P_R' represent promoters of lambda. t_1 is a transcription terminator.

A. Charon 34,35 linker

```
E        S  X S     C     X     S        P  B H     B
C        S  M M     L     B     A        S  S I     A
R        T  A A     A     A     L        T  P N     M
1        1  1 1     1     1     1        1  1 3     1

GAATTCGAGCTCGCCCGGGGATCGATCCTCTAGAGTCGACCTGCAGCCCAAGCTTGGATCC
----.----+----.----+----.----+----.----+----.----+----.----+-
1                                        50        61
```

ECR1,	1	SMA1,	16	HIN2,	37
TAQ1,	5	BIN1,	18	MNL1,	37
HNF3,	7	MBO1,	19	FNUH,	44
ALU1,	9	DPN1,	21	PST1,	45
BAN2,	11	TAQ1,	22	BSP1,	48
NSP2,	11	CLA1,	22	HIN3,	50
HGIA,	11	MBO1,	23	ALU1,	52
SST1,	11	DPN1,	25	BIN1,	55
XMA1,	14	BIN1,	28	BBV1,	55
AVA1,	14	XBA1,	29	MBO1,	56
HPA2,	15	MAE1,	30	BAM1,	56
SCR1,	15	HNF1,	33	XHO2,	56
NCI1,	15	SAL1,	35	DPN1,	58
SCR1,	16	TAQ1,	36	NLA4,	58
NCI1,	16	ACC1,	36	BIN1,	61

B. Charon 36 37,38 linker

```
E        S A   KXS     C     X     S        P  B  S H     B
C        S S   PMM     L     B     A        S  S  P I     A
R        T P   NAA     A     A     L        T  P  H N     M
1        1 1   111     1     1     1        1  1  1 3     1
                 /
GAATTCGAGCTCGGTACCCGGGGATCGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGATCC
----.----+----.----+----.----+----.----+----.----+----.----+----.---
1                                        50                  67
```

ECR1,	1	NCI1,	18	ACC1,	39
TAQ1,	5	SCR1,	19	HIN2,	40
HNF3,	7	NCI1,	19	MNL1,	40
ALU1,	9	SMA1,	19	PST1,	48
BAN2,	11	BIN1,	21	BSP1,	51
NSP2,	11	MBO1,	22	NLA3,	54
HGIA,	11	DPN1,	24	NSPC,	54
SST1,	11	TAQ1,	25	SPH1,	54
BAN1,	13	CLA1,	25	HIN3,	56
ASP1,	13	MBO1,	26	ALU1,	58
RSA1,	15	DPN1,	28	BIN1,	61
NLA4,	15	BIN1,	31	MBO1,	62
XMA1,	17	XBA1,	32	BAM1,	62
AVA1,	17	MAE1,	33	XHO2,	62

C. Charon 39, 40 linker

```
E        S A  KXS     C     X     S        P  B  S H     NX    A     S     X     N     N  B  N  B     A B              SB S     N
C        S S  PMM     L     B     A        S  S  P I     OM    V     P     H     H     C  G  A  B     P A              FG T     A
R        T P  NAA     A     A     L        T  P  H N     TA    R     E     O     E     O  L  R  E     A M              IL U     E
1        1 1  111     1     1     1        1  1  1 3     13    2     1     1     1     1  1  1  1     1 1              11 1     1
              /                                                 /                                                      /
GAATTCGAGCTCGGTACCCGGGGATCGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGCGGCCGCCCTAGGACTAGTCTCGAGGCTAGCCCATGGGGCGCCGGGCCCGGATCCATCGAGGCCAGGGAGGCCTGCCGGCGGGAATT
----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----+----.----
                                        50                                                  100                                                  149
```

ECR1,	1	KPN1,	17	BIN1,	31	HIN3,	56	MNL1,	77	NAR1,	101	NCI1,	110	ECR2,	124
TAQ1,	5	HPA2,	18	XBA1,	32	ALU1,	58	MAE1,	77	NLA4,	102	BAN2,	110	HAE3,	124
HNF3,	7	SCR1,	18	MAE1,	33	GDI2,	63	XHO1,	82	HHA1,	103	APA1,	110	HAE1,	124
ALU1,	9	NCI1,	18	HNF1,	36	GDI2,	63	AVA1,	82	HPA2,	104	NSP2,	110	SCR1,	126
BAN2,	11	SCR1,	19	SAL1,	38	XMA3,	63	TAQ1,	83	BBE1,	104	BIN1,	111	APY1,	126
NSP2,	11	NCI1,	19	TAQ1,	39	CFR1,	63	NHE1,	88	HAE2,	104	MBO1,	112	BGL1,	130
HGIA,	11	SMA1,	19	ACC1,	39	FNUH,	63	MAE1,	89	SCR1,	105	BAM1,	112	SFI1,	130
SST1,	11	BIN1,	21	HIN2,	40	NOT1,	63	NCO1,	94	NCI1,	105	XHO2,	112	HAE3,	133
BAN1,	13	MBO1,	22	MNL1,	40	HAE3,	65	STY1,	94	ASU1,	106	MNL1,	113	HAE1,	133
ASP1,	13	DPN1,	24	PST1,	48	FNUH,	66	NLA3,	98	ASU1,	107	DPN1,	114	STU1,	133
RSA1,	15	TAQ1,	25	BSP1,	51	STY1,	70	BGL1,	98	HAE3,	108	NLA4,	114	HPA2,	138
NLA4,	15	CLA1,	25	NLA3,	54	AVR2,	70	BAN1,	100	NLA4,	108	BIN1,	117	NAE1,	139
XMA1,	17	MBO1,	26	NSPC,	54	MAE1,	71	HINP,	101	HPA2,	110	TAQ1,	119		
AVA1,	17	DPN1,	28	SPH1,	54	SPE1,	76	AHA2,	101	SCR1,	110	MNL1,	122		

Figure 2

The linker sequences of all Charon vectors from Charon 34 to 40 are presented. The sequences were scanned for all class II enzyme recognition sites known as of December 1986. Those with six or more bases in the recognition sequence are presented above the sequence. All are tabulated below. The four letter abbreviations used for the restriction enzymes are as follows:

AAT2,Aat II/ACC1,Acc I/AFL2,Afl II/AFL3,Afl III/AHA2,Aha II/ALU1,Alu I/APA1,Apa I/APY1,Apy I,ASP1,Asp 718/ASU1,Asu I/ASU2,Asu II/AVA1,Ava I/ AVR2,Avr II/BAL1,Bal I/BAM1,BamH I/BAN1,Ban I(HgiC I)/BAN2,Ban II(Bvu I,HgiJ II)/BBE1,Bbe I/BCL1,Bcl I/BGL1,Bgl I/BGL2,Bgl II/BIN1,Bin I/BSP1,BspM 1/ BSP2,BspM II/BSS2,BssH II/BSTE,BstE II/BSTX,BstX I/CFR1,Cfr I(Eae I)/CFR2,Cfr 10I/CLA1,Cla I/DPN1,Dpn I/DRA1,Dra I/DRA3,Dra III/ECR1 EcoR I/ECR2,EcoR II (BstGII,BstNI)/ECRV EcoR V/ESP1,Esp I/FNUH Fnu4H I/GDI2, Gdi II/HAE1,Hae I/HAE2,Hae II/HAE3,Hae III/HGIA,HgiA I/HHA1,Hha I/HIN2,HinD II/HIN3,HinD III/ HINP,HinP I(HhaI)/HNF1,HinF I/HNF3,HinF III/HPA1,Hap I/HPA2,Hpa II/KPN1,Kpn I/MAE1,Mae I/MBO1,Mbo I/MLU1,Mlu I/MNL1,Mnl I/MST2,Mst II/NAE1,Nae I/ NAR1,Nar I/NCI1,Nci I/NCO1,Nco I/NDE1,Nde I/NHE1,Nhe I/NLA3,Nla III/NLA4,Nla IV/NOT1,Not I/NRU1,Nru I/NSI1,Nsi I/NSP2,Nsp II(Bsp1286)/NSPC,Nsp HI/ PPU1,PpuM I/PST1,Pst I/PVU1,Pvu I(Xor I)/PVU2,Pvu II/RSA1,Rsa I/RSR2,Rsr II/SAL1,Sal I/SCA1,Sca I/SCR1,Scr FI/SFI1,Sfi I/SMA1,Sma I/SNAB,Sna BI/ SPE1,Spe I/SPH1,Sph I/SSP1,Ssp I/SST1,Sst I/SST2,Sst II/STU1,Stu I(Aat I)/STY1,Sty I/TAQ1,Tag I(TthB81)/TTH1,Tth 111I/XBA1,Xba I/XHO1,Xho I(Ccr I)/ XHO2,Xho II/XMA1,Xma I/XMA3,Xma III/XMN1,Xmn I(Asp700)

RESULTS

RATIONALE

In general, replacement vector construction involves removal of restriction sites from essential genes of the vector, removal of nonessential segments of DNA from the vector arms, installation of polylinker cloning sites, and provision of stuffer DNA to reserve space for insertion of foreign DNA. The main desirable features of the vector arms are minimum size (high DNA capacity), multiplicity of cloning sites and ability to grow in recombination deficient conditions. Ease of removal and the ability to detect inadvertant reinsertion are the main features of desirable stuffers. To increase versatility and to minimize the number of different vectors needed for cloning with different restriction enzymes, we elected to successively combine desirable features in a single vector rather than to create different vectors for different purposes. Charon 40 thus combines most of the desirable features of all previous replacement vectors.

The starting point for the new vectors was Charons 34 and 35, (6) which are identical to each other except for the stuffer fragments (see Figure 1.) These vectors had been designed for cloning with six restriction enzymes (EcoR I, Xba I, Sst I, Sal I, Hind III, and BamH I) under recombination deficient conditions.

The first new vector, Charon 36 features the ability to clone with Kpn I, an enzyme that cuts eukaryotic DNA infrequently. The second, Charon 37 was further enhanced by addition of the ability to clone with Sma I, a flush cutting enzyme, and Xma I, its isoschizomer, that leaves a 5′ overhang. Sma I is particularly useful because it can be used to clone blunt-ended fragments which can be produced by DNase I digestion or shear after filling in the ragged ends. In addition, about 1 kbp of increased cloning capacity was achieved by modification of the right arm of Charon 37.

The main objectives for Charon 38 were addition of another kbp of capacity by shortening of the left arm and removal of the Apa I site in the essential lambda gene G so that Apa I cloning would be possible in the next vector of the series. Charon 39 incorporated a new polylinker which added eight more cloning enzymes (Xma III, Not I, Avr II, Spe I, Xho I, Apa I, Sfi I and Nae I). To design the polylinker we scanned the DNA sequence of the Charon 39 vector arms by computer to identify sites that were absent or present once and which lacked ambiguties in the overhangs produced by digestion. The chosen recognition sites were appended to the polylinker sequence of Charon 37 beginning at the Hind III site in more or less random order except that we avoided adjacencies that created more than one instance of any site in the polylinker. Bam H1, Sfi I, and Nae I were placed at the interior end; Sfi I is useful for detachment of the stuffer in one piece and

Nae I for cutting it up. We wished to retain the Bam H1 site at the interior so that Sau 3A partial clones could always be cut out with the maximum number of enzymes. We did not attempt to shorten the sequence by overlapping recognition sequences because it was felt that in some cases two enzymes might need to be used to cut the same DNA. Sequences of all of the polylinkers are provided in Fig. 2.

Charon 39 also incorporated a new type of stuffer fragment that we term polystuffer. The basic idea of the polystuffer is to provide a segment of DNA in the disposable region of the vector with very closely spaced restriction sites so digestion will produce small fragments. These can be separated from the vector arms simply and with little loss of DNA by polyethylene glycol PEG precipitation or by other methods that exploit the size difference. The simplest way to accomplish this is to create the stuffer as a repetitive DNA segment where the repeat includes the restriction site to be used for polystuffer degradation. To create the polystuffer we devised a method for polymerization of small DNA fragments that yields only direct repeats since perfect inverted repeats are quite unstable in recombinant proficient *E. coli* (15,16), (see methods). A refinement of the polystuffer is inclusion of a detectable marker on the repeat to allow confirmation by the investigator that no copies of the stuffer have been inadvertantly reincorporated into his clone.

The rationale for the design of the repeat unit of the Charon 39 polystuffer was as follows: we desired a directionally polymerizable fragment with a Nae I site and a lac operator. The lac operator serves as a stuffer indicator since even a single copy within a phage can induce chromogenically detectable β-galactosidase in lac+ hosts by means of repressor titration (8). We wished to include Nae I because blunt-ended fragments are more difficult to reclone and are easier to separate by PEG precipitation than are those with single strand overhangs. Nae I is one of the only commercially available flush cutting enzymes that does not cut λ. We chose an adenovirus segment as the source of the Nae I site because its Nae I sites are known to cut relatively well (33). Considerable unexplained variation in the specific rate of site cleavage has been observed for Nae I depending on flanking sequences around the recognition site (33). Moreover, in this first attempt we wanted to be able to clone the lac operator a considerable distance away from Nae I, to minimize the possibility of decreasing its site cleavage rate as a result of inserting the operator sequence. The lac operator was inserted at the BamH I site, 83 bp away from Nae I. The polystuffer of Ch 39 consists of a 540 bp lac operator containing fragment present in the vector in about 25 tandem copies.

The Charon 39 polystuffer repeat unit has six cloning enzyme sites in

Table I

Vector	Arm Lengths		Capacities *		Cloning Sites
	Left arm	Right Arm	Min	Max	
Charon 35	19470	10615	7915	22915	EcoRI,XbaI,SstI,SalI,HindIII,BamHI
Charon 36	19482	10615	7903	22903	add KpnI
Charon 37	19482	9622	8896	23896	add SmaI
Charon 38	19171	9622	9207	24207	Same sites as Charon 37
Charon 39	19171	9622	9207	24207	add XmaIII,NotI,AvrII,SpeI, XhoI,ApaI,SfiI,NaeI
Charon 40	19171	9622	9207	24207	Same sites as Charon 39
Ch 39, 40	19171	8939	9890	24890	Nhe I++

* Assumes 38000 (min) and 53000 (max) basepairs for lambda packaging limits
+ The min-max capacities of Charons 35 and 36 for Sal I fragments are the same as Charon 37
++Nhe I clones are gam⁻. Nhe I, Xba I, Spe I, and Avr II fragments are all compatible with Nhe I phage arms.

addition to Nae I. After cutting with any of the cloning enzymes in the repeat, the stuffer can be separated very conveniently from the arms by precipitation (see methods). For cloning enzymes that do not cut in the polystuffer repeat, Nae I can be used in combination with the cloning enzyme to degrade the stuffer. Since Nae I is the most interior site in the polylinker this enzyme not only cuts up the repeat unit but completly detaches the polystuffer from the polylinker.

Nae I should probably not be used in combination with a cloning enzyme that also cuts within the polystuffer. In this situation, the use of Nae I results in fragments without the lac operator that cannot be chromogenically detected. Moreover, these have ends corresponding to the vector arms that make it likely they will be recloned if they are not successfully removed.

The Charon 40 polystuffer was designed to place the lac operator very close to the Nae I site so the stuffer repeat could be shorter despite the fear that cleavage at Nae I might be interfered with. This fear proved unfounded and the polystuffer of Charon 40 consists of about 80 copies of a shorter repeat of 235 bp. Only Nae I and Xma III cut in the polystuffer repeat. Thus it is practical to cut the vector with Nae I in all cases except when Xma III is the cloning enzyme. This shorter DNA segment is also more easily removed with polyethylene glycol.

Three of the new vectors are available as "A" versions. For Charon phages the A suffix indicates the presense of amber mutations in essential genes necessitating a supressor strain for propagation. Three of the new vectors, Charons 38-40 were constructed in both wild type and A versions. Double amber mutations Aam 32 and Bam 1 were used to ensure low reversion rates. Charons 38-40 A are useful for implementing several cloning strategies that use a supressor tRNA gene to "tag" the gene of interest for

cloning (10,18). Another use of the ambers is to provide biological containment.

Library construction and cloning efficiencies

The new vectors can be used to generate libraries of large genomic DNAs cut with any of the cloning enzymes listed in Table I. Using Ch 40 carms prepared via the rapid PEG precipitation protocol and "target" genomic DNA fragments prepared by conventional methods (13), we have obtained cloning efficiencies of 2-4.5 x 10^6 plaques/ug of insert at minimum.

VECTOR CONSTRUCTION

Construction of Charon 36

The chief goal for Ch 36 was provision of the ability to clone with Kpn I. The prerequisites are removal of the two Kpn I sites from the essential J gene in the lambda left arm and provision of polylinkers containing Kpn I sites. The presence of two sites renders the "cycling" method of mutant selection extremely difficult due to the statistical rarity of double mutants in 12 base pairs of DNA. A site specific mutagenesis strategy was thus adopted. A synthetic oligonucleotide duplex, Kpoligo, was synthesized (See Fig. 3.) It has Kpn I-compatible 3' overhangs but its sequence is such that it destroys Kpn I sites after ligation into them. We hoped that the three amino acid insertions (serine or glycine depending on the orientation) into the J gene would be compatable with its function as a tail fiber. Since Charons 34 and 35 both possess multiple Kpn I sites in their stuffers, a clone, Charon 35-H668, isolated from a human Ch35 Hind III library which lacked Kpn I sites in the insert was used for the mutagenesis precursor. This phage was digested to completion with Kpn I and ligated with the oligonucleotide duplex in a 1:1 molar ratio. The ligation products were recut with Kpn I to select against wild type recombinants and packaged *in vitro*. All recombinants tested lacked the leftmost Kpn I site of the J gene but not the other, implying that the three amino acid insertion at the right hand position is not tolerated. However, the removal of one Kpn I site allowed us to readily remove the other by the conventional cycling method (7) to produce Ch 35-H668 Kpn^-.

The new vector also required a polylinker sequence containing a Kpn I site. This was obtained from pUC19 in exactly the same way that the Ch 35 linker had been derived from pUC13, (see Fig 1 of reference 3). Specifically, a plasmid lacking the BamH I site of pUC19, pUC19Bam- was prepared by filling in BamH I linearized pUC19 with the Klenow fragment of DNA polymerase I and recircularizing the product. This plasmid was then linearized with Hind III and ligated to Hind III-cut pIN-II-AI (19). After redigestion with BamH I, the linker was purified from an acrylamide gel and ligated to

STRUCTURE OF VECTORS

Ch34R = LAM (31753,33248) + (34500,35847) + (37926,39167) + X + LAM (39169,40501) + (43308,44140) + X + LAM (44142,44971) + X + LAM (44973,48502)

Ch34L = LAM (1,5504) + X + LAM (5506,19368) + LACZ (4409,4308)

Ch36L = LAM (1,5504) + X + LAM (5506,17057) + KPOLIGO (1,8) + LAM (17054,18556) + X + LAM (18558,19368) + LACZ (4409,4308)

Ch37R = SALR1 (10<7) + LAM (32746,33248) + (34500,35847) + (37926,39168) + X + LAM (39170,39888) + X + LAM (39890,40501) + (43308,44140) + X
+ LAM (44142,44971) + X + LAM (44973,48502)

Ch38L = LAM (1,5504) + X + LAM (5506,10090) + X + LAM (10092,17057) + KPOLIGO (1,8) + LAM (17054,18556) + X + LAM (18558,19159) + SALR1 (2,9)

Ch34 = C34L + C34 LINK (1,61) + C34 STUFF + C34 LINK (61<1) + C34R

Ch35 = C34L + C34 LINK (1,61) + C35 STUFF + C34 LINK (61<1) + C34R

Ch36 = C36L + C36 LINK (1,67) + C34 STUFF + C36 LINK (67<1) + C34R

Ch37 = C36L + C36 LINK (1,67) + C34 STUFF + C36 LINK (67<1) + C37R

Ch38 = C38L + C36 LINK (1,67) + C34 STUFF + C36 LINK (67<1) + C37R

Ch39 = C38L + C39 LINK (1,149) + [pUC19 (421<402) + AD2 (21338,21610) + LACOP (1,25) + AD2 (21607,21816)] + GATC + C39 LINK (149<1) + C37R

Ch40 = C38L + C39 LINK (1,149) + [AD2 (21607,21724) + LACOP (1,25) + AD2 (21725,21816)] + GATC + C39 LINK (149<1) + C37R

```
KPOLIGO:        1     8          SALR1:    1          15          LACOP:  1                    25
                TGGAGGAAGTAC                 GTCGAGGGGGAATTC                  GGAATTGTGAGCGGATAACAATTCC
           CATGACCTCCTT
```

Fig. 3: DNASTAR notation (9) is used to precisely describe the sequence of each vector by itemizing the exact endpoints of subregions of sequences that can be joined together to produce the vector sequence. Numbers in parentheses separated by a comma refer to segment endpoints (inclusive numbering). The left arrow refers to the reverse complement of the sequence from the first to the second numbered position. Definition of the right and left arms of Ch 34 (C34R and C34L respectively), the left arm of Ch 36 (C36L), the right arm of Ch 37 (C37R) and the left arm of Ch 38 (C38L) are sufficient to allow all the new vector genomes to be compiled using published sequences plus the linker sequences (Fig 2) and this figure. GATC refers to four bases joining the polystuffer with the polylinker. X refers to an unknown residue replacing one base of each restriction site removed by mutation. C34 STUFF and C35 STUFF refer to unsequenced large stuffer fragments. The orientation of KPOLIGO and LACOP in these constructs was not determined. Polystuffer sequences can be derived by repetition of the sequences within the square brackets for Ch 39 and Ch 40. References: LAM (29), LACZ (30), AD2 (Adenovirus-2, 31), pUC19 (32).

the BamH I stuffer fragment of Ch 34 which lacks internal BamH I and EcoR I sites. The preparation was then recut with EcoR I, purified on a gel and ligated with the EcoR I arms of the Kpn- mutant phage. Resulting phage were screened by probing with 5'-labelled EcoR I - Sma I linker fragment from pUC19, which hybridizes to the pUC 19 linker and not to the pUC 13 linker under the conditions used (see methods). Positively-hybridizing phage were purified and shown to have the desired structure by restriction mapping. The new phage retains all of the features of Ch34 and Ch35 with the addition of Kpn I-cloning capacity.

Construction of Charon 37

To introduce the ability to clone with Sma I, (and Xma I), the Sma I site in the P gene of lambda on the right arm had to be removed. Since Charons 34 to 36 all have additional Sma I sites in their stuffers or polylinkers, we used the cycling procedure on Charon 33, which has the same right arm but no extra Sma I sites (6). The resulting phage, Charon 33 Sma^-, was fully characterized by restriction mapping. To produce a phage with a shorter right arm we chose the Sal I site immediately to the left of the gam gene for attachment of the polylinker and stuffer. This eliminated non-essential DNA (beta and part of exo) in between the Sal I site and the EcoR I site that had been used in previous vectors Charons 33 to 36. Since the only Sal I sites needed for cloning are in the polylinkers, it was necessary to connect the Eco R1 end of the polylinker to the Sal I site of the right arm in such a way as to preserve the Eco R1 site while eliminating the Sal I site at the junction. This was done by constructing an EcoR I-Sal I "adapting" plasmid (pUC8Sm/Xh) containing a short Xho I - EcoR I fragment. Because the product of ligation of an Xho I site with a Sal I site is a sequence that can be cut by neither enzyme, the Sal I site could be eliminated. The plasmid was constructed by inserting a flush ended Xho I linker into the Sma I site of pUC8 (20). Ch 33 Sma^- DNA cut with Sal I was ligated to adapter plasmid DNA cut with Xho I. Sal R1 (Fig 3) refers to the junction sequence resulting from this ligation. The preparation was then cut with EcoR I and purified on a gel to obtain a short right arm with an EcoR I site at the end and no Sma I, Sal I or Xho I sites. The preparation was then cut with EcoR I and the right arm isolated. Next a Ch 36 preparation was cut with EcoR I, ligated with the new right arm and packaged. Resulting plaques were screened for failure to hybridize with labeled 1054 bp EcoR I-Sal I fragment from Ch 36 which should be absent from the new phage construct. Charon 37 was shown to have the appropriate structure by extensive restriction mapping.

Construction of Charons 38 and 38A

The starting point for this construction was λFAC101, a clone made in

Charon 30 A by Francine Lawyer and David Gelfand of Cetus Corp. As a result of the cloning methodology used in its construction (see methods), λFAC101 has a Sal I site at position 19159 where an Alu I site had been in wild type lambda. This is immediately to the right of gene J, the last essential gene of the left arm, so the lac 5 remnant and part of the truncated lom gene in previous Charon phages can be eliminated producing a minimal length left arm.

The first step was removal of the single Apa I site in gene G of λFAC101 by cycling so that this enzyme could be used for cloning. Next, to retain the ability to clone with Kpn I it was necessary again to remove the two Kpn I sites in gene J. This was done by electrophoretically isolating the 2.4 kbp Sca I fragment from Ch 36 (co-ordinates 16420 to 18683) which encompasses the two mutant Kpn I sites. λFAC 101 was also cut with Sca I and all the fragments except the 2.4 Kbp one were isolated from the gel. There are other Sca I sites in this phage but none which also cut within essential genes. These fragments were ligated together and packaged. Since an intact J gene is essential for growth, the desired recombinant constituted the majority of viable phages isolated from this experiment. The structure was confirmed by digestion of candidates with Kpn I, Apa I and Sal I.

The Sal I left arm from λFAC 101 $Apa^{-}Kpn^{-}$ was connected to the poly-linker, stuffer and right arm of Charon 37 using the same adapter plasmid strategy as had been used to attach the right arm of Charon 37 in the previous construction.

The resultant recombinant, designated Charon 38A, was shown to have the appropriate restriction sites. The presense of amber mutations was confirmed by the failure of Ch 38A to produce even one plaque when 5×10^9 phages were plated on suppressor negative strains (25 plates). This low reversion rate indicates that both ambers are present but they were not tested by complementation.

The next step was isolation of Charon 38 by reversion of the amber mutations of Charon 38A. The low reversion rate of Ch 38A on supressor negative bacteria made a direct approach technically difficult. We observed that Ch 38A is poorly supressed on an Su II host. The plaques are small and liquid lysate titres are an order of magnitude lower than on the Su II, Su III host normally used. Thus spontaneous amber revertants at one locus appearing within the phage population should have a significant growth advantage over the parentals, allowing the accumulation of double revertants. After two growth passages in K802 (Su II) liquid cultures, the reversion rate of Ch 38A measured on the supressor negative strain had fallen to 10^{-6} and Ch 38 candidates could be picked directly from these plates.

Construction of Charons 39 and 39A

Construction of Charon 39 required both the extended polylinker and the polystuffer. To generate the polystuffer we prepared a plasmid containing one copy of the polystuffer repeat, a 540 bp BamH I/Bgl II fragment. This contained a part of the pUC19 linker (including BamH I, EcoR I, Kpn I, and Sma I sites) and adenovirus-2 DNA from coordinates 21338 to 21816. This sequence was manipulated so that the internal BamH I site (adenovirus position 21606) was removed and a synthetic lac operator (Pharmacia, Fig.3) was inserted. To do this, an EcoR I fragment from adenovirus-2 (coordinates 21338 to 25633) was cloned into the EcoR I site of pUC19. This adenovirus piece included the rapidly cutting Nae I site flanked by BamH I and Bgl II sites. The EcoR I clone was cut with BamH I, rendered blunt-ended with Pol I/dNTPs, and ligated with a large excess of blunt-ended non-phosphorylated lac operator DNA. Next the product was ligated with an excess of phosphorylated synthetic BamH I linkers. Subsequent digestion of the preparation with BamH I and Bgl II allowed the isolation of a 540 bp fragment. This was cloned between the BamH I/Bgl II sites of plasmid vector pKC30 (21,22). Clones containing the lac operator sequence were selected on a lac+ host (K802) by their ability to induce β-galactosidase by the repressor titration effect. The plasmid restriction map verified the ability to remove the fragment with BamH I and Bgl II, and confirmed the presence of the other expected sites. The sequence of this polystuffer unit can be compiled from the listed segments in Figure 3., except in the orientation of the lac operator segment, which was not determined.

Next the purified 540 basepair BamH I/Bgl II fragment isolated in quantity from the plasmid was directionally polymerized as described in Methods. Directionalization was demonstrated by the complete resistance of the product to BamH I or Bgl II digestion, in contrast to its sensitivity to Xho II attack at BamH I/Bgl II junction sequences. Xho II re-digestion was used to obtain polymer in the desired size range (10-20 kb). Gratifyingly, the polymer was completely and rapidly digested with Nae I. This preparation was purified, blunt-ended with Pol I/dNTPs and attached to the polylinker described below.

The polylinker for Charon 39 was synthesized as two double stranded oligonucleotides with appropriate overhangs. These were ligated together and cloned between the Hind III and Sma I sites of pUC8. In this plasmid (pISD27) a Hind III- EcoR I fragment contains the new part of the polylinker which had to be attached to the polylinker of Ch 38 and to the polystuffer. pISD27 DNA was cut with EcoR I and Pol I was used to fill in the sticky ends. This preparation was ligated to the polystuffer, recut with Hind III and repurified. This was then ligated into Hind III cut arms of Ch 38A.

Packaged phage were plated on ED8767 in the presence of chromogenic β-galactosidase substrate, and resulting intensely blue plaques were selected. A clone with all of the expected features was isolated and termed Ch 39A. Charon 39 was derived from a conventional genetic cross between Charon 38 and 39A selecting for blue plaques on a lac^+ Su0 - strain.

Construction of Charon 40 and Charon 40A

The goal for Charon 40 was to provide a better polystuffer than the one in Charon 39, i.e. a shorter segment with as few sites as possible other than Nae I. For this purpose a smaller adenovirus fragment BamH I/Bgl II (21607-21820) was chosen. To clone this, a new BamH I - Bgl II vector, pISD34, was constructed by inserting a Bgl II linker into pBR322 at the Xma III site. A new vector was needed because pKC30 has an internal Xma III site that would interfere with subsequent steps. The adenoviral clone retained both BamH I and Bgl II sites and had a unique Xma III site close to the Nae I site. This was cut, filled in and then the synthetic lac operator fragment was inserted. Regrettably, this regenerated Xma III sites on either side of the operator sequence. Clones were selected in the same manner as for the 540 bp fragment described above. Although as noted above an insertion close to the Nae I site had the potential for adversely affecting its cleavage rate, we did not observe any significant problems in this regard with this clone. This 235 bp "tagged" repeat unit was treated in an exactly equivalent manner as for the Ch 39 540 bp fragment to produce Ch 40A.

Charon 40 was produced by cutting Charon 40A and Charon 38 with Eco R1, ligating them together and packaging. The same selection as for Charon 39 was used to identify the Charon 40 clone lacking an amber mutation.

DISCUSSION

Table 1 summarizes the predicted DNA capacities of the new vectors. Their primary application is in the cloning of large DNA fragments with a large number of cloning enzymes in the size range 8 to 24 kbp. With Charons 39 and 40, 16 cloning sites are available in the polylinker. In addition a 17th enzyme, Nhe I, can be used to clone fragments of almost 25 Kbp but the gam gene and the right hand copy of the polylinker are both removed. Thus they cannot be propagated on recA$^-$ but they could be grown on recA-recBC$^-$ strains. All the new vectors produced lysate titers of 1 x 10^{10} or greater on the standard LE 392 host with the liquid lysate protocol and media described in methods. We could not see a significant difference in yield between these vectors and EMBL 3 (3) under these conditions. We did observe that Charons 36 to 40 are relatively unstable in the absense of magnesium, evidently as a consequence of the mutations in the J gene used to remove Kpn I sites. Provided Magnesium ions are maintained at 10mM (as is the case for

RESTRICTION SITES OF CHARON 40

Left Arm		BAL1:	6498	PVU2:	12101	Polylinker		NCO1:	19766	PVU1:	21762
LEND:	1	DRA3:	6640	PVU2:	12164	ECR1:	19180	Polystuffer		BGL2:	21997
XMN1:	37	ECRV:	6683	SNAB:	12190	SST1:	19190			BSTX:	22193
DRA1:	92	BSTX:	6713	STU1:	12436	ASP1:	19192	Polylinker		NSI1:	22205
PVU2:	211	BAL1:	6879	BSP2:	13086	XMA1:	19196	NAE1:	19820	NDE1:	22252
BGL2:	415	BSP2:	6894	XMN1:	13106	KPN1:	19196	STU1:	19826	BGL2:	22648
MLU1:	458	BSTE:	7058	BSTX:	13270	SMA1:	19198	SFI1:	19832	BGL2:	22708
BSP2:	610	BAL1:	7586	BSTE:	13348	CLA1:	19204	BAM1:	19843	DRA1:	22729
ECRV:	652	BSP2:	7811	ECRV:	13437	XBA1:	19211	APA1:	19853	SSP1:	22870
HPA1:	734	PVU2:	7833	BSTE:	13572	SAL1:	19217	NAR1:	19856	ECRV:	23248
XMN1:	1155	HPA1:	7950	BSTE:	13689	SPH1:	19233	BBE1:	19859	SPH1:	23316
BAL1:	1328	BAL1:	7980	BSP2:	13806	HIN3:	19235	NCO1:	19861	ESP1:	23345
BSP2:	1826	BAL1:	8058	BCL1:	13820	XMA3:	19242	NHE1:	19867	HPA1:	23502
PVU2:	1919	ECRV:	8086	BAL1:	13936	NOT1:	19242	XHO1:	19873	HPA1:	23730
ECRV:	2086	HPA1:	8201	RSR2:	13984	AVR2:	19249	SPE1:	19879	STU1:	23888
BAL1:	2208	BSTE:	8322	ECRV:	14025	SPE1:	19255	AVR2:	19885	BSP2:	23895
SPH1:	2216	BSTX:	8420	BSTX:	14345	XHO1:	19261	XMA3:	19892	BSTE:	23943
XMN1:	2323	DRA1:	8462	DRA3:	14482	NHE1:	19267	NOT1:	19892	NDE1:	24026
PVU2:	2387	SSP1:	8471	BSS2:	14815	NCO1:	19273	HIN3:	19899	SST2:	24283
PVU2:	2528	XMN1:	8494	BAL1:	14905	NAR1:	19280	SPH1:	19909	BCL1:	24770
BSP2:	2619	ECRV:	8824	AAT2:	14978	BBE1:	19283	SAL1:	19917	CLA1:	24913
PPU1:	2816	BCL1:	8844	HPA1:	14993	APA1:	19289	XBA1:	19923	NCO1:	25336
BSTX:	2862	BSTX:	8857	MLU1:	15372	BAM1:	19291	CLA1:	19932	XMN1:	25819
DRA3:	2959	BAL1:	8861	CLA1:	15584	SFI1:	19309	XMA1:	19938	AAT2:	26655
PVU2:	3060	DRA3:	9004	BSTE:	16012	STU1:	19312	SMA1:	19940	AAT2:	26684
BAL1:	3262	BSTE:	9024	BSP2:	16040	NAE1:	19318	ASP1:	19942	NAR1:	26768
BSP2:	3329	BCL1:	9361	PVU2:	16080	Polylinker		KPN1:	19946	BBE1:	26771
BSP2:	3341	AAT2:	9398	CLA1:	16121			SST1:	19952	XMN1:	26833
BSS2:	3522	BSP2:	9696	DRA1:	16296	Polystuffer		ECR1:	19954	ECRV:	26916
PVU2:	3639	XMN1:	10115	SCA1:	16423	NCO1:	19333	Polylinker		BCL1:	27454
RSR2:	3801	ESP1:	10298	ESP1:	16519	TTH1:	19380			CLA1:	27527
BSS2:	4126	BSP2:	10318	BSS2:	16649	SST2:	19406	Right Arm		BSTX:	27529
BAL1:	4195	NSI1:	10329	XMN1:	16913	XMA3:	19444	SCA1:	20027	SSP1:	27935
CLA1:	4199	BAL1:	10611	BSP2:	17642	XMA3:	19473	CLA1:	20187	DRA3:	28405
NRU1:	4592	ESP1:	10683	ECRV:	17781	NAE1:	19477	NSI1:	20194	SSP1:	28411
BSP2:	5098	BAL1:	10779	MLU1:	17803	PVU2:	19520	STU1:	20222	SSP1:	28511
AAT2:	5109	BSTX:	10922	BSTE:	17953	NCO1:	19527	BSTX:	20575	DRA1:	28519
HPA1:	5269	TTH1:	11205	BSTX:	18048	NCO1:	19572	NHE1:	20651	SSP1:	28646
MLU1:	5548	AAT2:	11247	ASU2:	18061	TTH1:	19619	CLA1:	20669	XMN1:	28656
DRA3:	5618	HPA1:	11585	ECRV:	18399	SST2:	19645	SSP1:	20975	SSP1:	28783
BSS2:	5627	ESP1:	11662	SCA1:	18698	XMA3:	19683	CLA1:	21023	BCL1:	29030
BSTE:	5687	PVU1:	11936	Left Arm		XMA3:	19712	HPA1:	21233	DRA3:	29527
HPA1:	5710	BSP2:	11986			NAE1:	19716	BSP2:	21421	PPU1:	29562
RSR2:	6042	SPH1:	12006			PVU2:	19759	BGL2:	21683	REND:	29590

Fig. 4: The enzymes that cut Charon 40 fewer than 20 times are presented in the order they cut. The sequence as scanned by the computer has two copies of the polystuffer. Enzyme abbreviations are as in Fig. 2. LEND and REND refer to the same left end and right end. The same diagram applies to Charon 39 except for the polystuffer.

recommended GIBCO phage media), no problems have been encountered. Yields on the recA$^-$ strain K802recA were consistently lower by a factor of 5 to 10. We therefore suggest that the recA strain be used only when necessary. Gam expression from the vectors is expected to substantially reduce recombination by inhibition of the recBC function even on rec$^+$ strains. But the polystuffer size certainly varies on LE392 as indicated by electrophoresis on agarose gels. This causes no problem in growing phage for making vector arms but it shows that recombination is not low enough on that strain to propagate highly repetitive DNA. We suggest the use of recA$^-$ hosts in such cases.

Several approaches were developed in the past to detect recombinants ameliorating the problem of inadvertant recloning of the replacable "stuffer

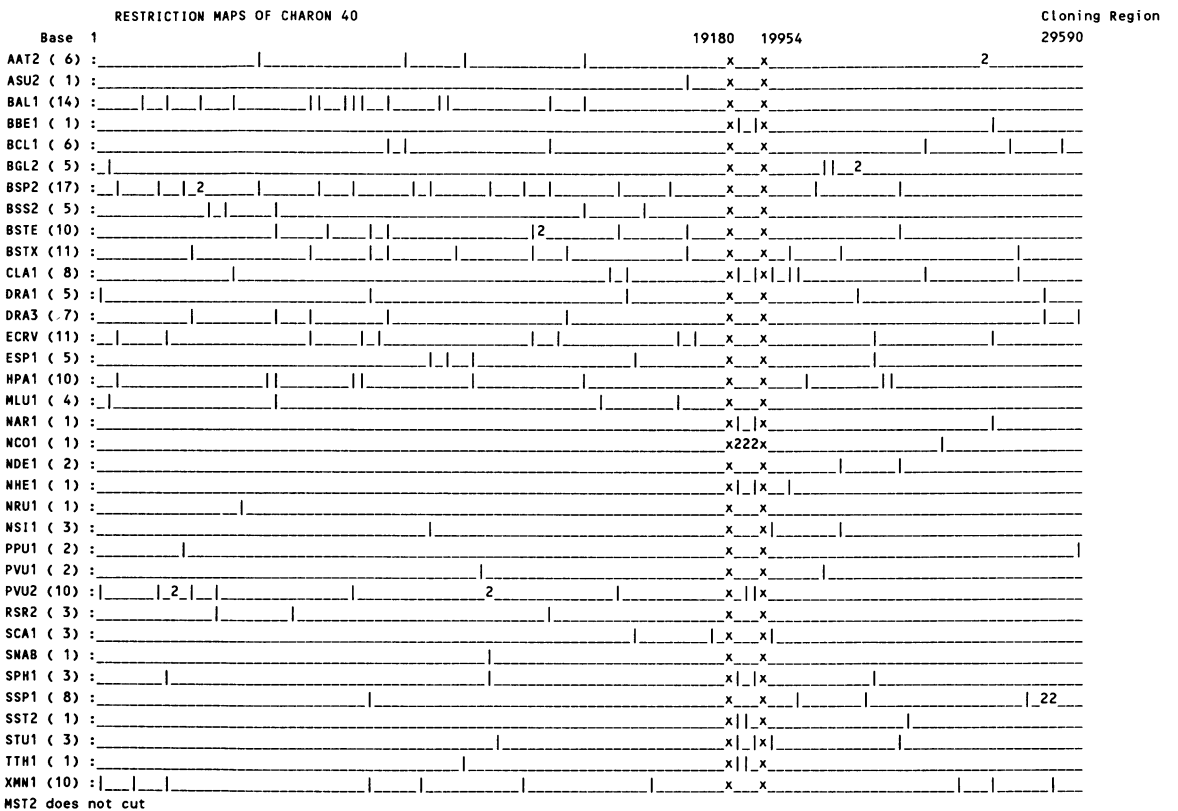

Fig. 5: The enzymes that cut the Charon 40 vector arms fewer than 20 times are presented as mini maps. The sequence scanned by computer had two copies of the polystuffer. The four character enzyme name abbreviations (see Fig. 2) are followed by the number of cuts in parentheses (arms only). Cut sites are indicated at their approximate positions. "2" refers to double cuts. "X"s define the ends of the arms. Base positions of the ends of the arms are listed across the top. The same diagram applies to Charon 39 except for the polystuffer.

fragment". These include physical removal of the stuffer fragment on gels or gradients (8), Spi selection on P2 lysogens (23) or colored plaques for screening against it (8) and the use of alternate cutting restriction enzymes to reduce the clonability of the stuffer fragment (24).

The EMBL series of phages (3) allow a method that is somewhat akin to our polystuffer approach. Sfi I sites are located very close to the two ends of the stuffer. Propyl-alcohol precipitation can thus be used to remove the stuffer end pieces so that the remaining fragments have inappropriate termini for cloning. Thus the stuffer is theoretically excluded from ligation with the vector arms even though it is not physically removed, whereas appropriately-cut target DNA can be cloned normally. Although this procedure eliminates the necessity to remove the large stuffer fragment from ligations with target DNA, total removal of the stuffer is probably preferable. A primary requirement for either the EMBL scheme or the polystuffer approach is that the stuffer lack all cloning enzyme sites. This is difficult to achieve with conventional stuffers as more and more cloning enzymes are added. Although sites can of course be exhaustively removed *in vitro* we chose to construct an artificial polymeric stuffer with few sites which was designed to be readily removable, easily detectable and totally defined as to its sequence.

The polystuffer technique has been a great practical boon for the generation of libraries. Separation of arms by the PEG procedure as described in Methods is fast and easily carried out. Recovery of DNA is much higher than with gels or gradients so less DNA from the vector must be grown. Purified arms produced in this way are also very efficient for cloning. There is no need for special selection for recombinants, for example, on P2 lysogens. Removal of the stuffer is virtually complete so that all arms are available to react with target DNA. Almost no plaques in these libraries fail to incorporate target DNA and stuffer reincorporation as assayed chromogenically is less than 0.1% percent in typical cases.

All the vector portions of these phages came from sequenced sources. Figure 3 shows the exact endpoint of sequences that were combined to create computer sequence files for each vector. There are several positions in the sequence where restriction sites have been eliminated by unknown, presumably single-base changes. These are indicated with Xs. We are also uncertain as to the orientation of the Kpn I oligo in the J gene and the lac operator in the two polystuffers. With the exception of possible effects caused by these few mutations, restriction patterns can be predicted by computer scan. The Charon 40 file was scanned for known recognition site sequences, and enzymes that cut fewer than 20 times (assuming two copies of the polystuffer) were listed. Figs. 4 and 5 show this complete restriction map as predicted from

the sequence. With the aid of these figures, users can easily map the restriction sites in Charon 40 clones against the background of vector fragments.

In the A series of vectors, amber mutations can be used for the recombinational-screening procedure of Seed (25), or as vectors for the selective isolation from genomic DNA preparations of sequences flanking an inserted suppressor tRNA gene. Retroviral vectors bearing the Sup F gene have been described (17,18), which, in a manner analogous to transposon tagging, can be used to clone eukaryotic sequences adjacent to the point of insertion. Genomic cloning of DNA from cells infected with such vectors should be possible with our new phages. The amber mutations should also provide biological containment.

Cosmid vectors carrying spacer fragments polymerized in a head-to-tail fashion have been recently described (26). The polymeric segments of such vectors were designed as spacers to allow the cloning of smaller DNA fragments when with conventional cosmids, by rendering clones large enough to be packaged *in vitro*. They differ fundamentally from the new polystuffer phage in that they are insertion vectors; hence the cosmid "polystuffers" are necessarily retained in cloned species.

Further modifications of λ vectors to increase capacity are possible, but it is becoming increasingly difficult to make improvements. It has been reported recently that 2000 basepairs of λ DNA in the right arm between the Rz gene and cosR is non-essential (27). About 500 basepairs of nonessential DNA between gene N and gam also exists. Thus deletions could be engineered in the present Charon series to increase their capacity by 2500 basepairs. By eliminating genes gam, N and Rz which are nonessential for growth but which do confer advantages on the vector, an additional 1000 basepairs could be achieved. That would seem to be the ultimate capacity limit for λ replacement vectors. It is interesting to speculate that such a deleted vector might package two to the capsid, thereby eliminating a need for a stuffer at all. A phage that has two equally spaced cos sites has been described and packages well (28).

ACKNOWLEDGEMENTS

We thank David Gelfand, Richard Malavarka, Francine Lawyer, Bruce Kaplan, John Rossi, Val Darby, Pim Stemmer, Nola Peterson, Anna Shen, Patricia Wathen and Andrew Hopfensperger for help; Support from Leukemia Society, Gibco Inc., Merck Inc.; NIH grant GM21812. This is paper No. 2915 from the Laboratory of Genetics, University of Wisconsin, Madison, WI.

[1]Present address: CSIRO Div. of Mol.Biol. PO box 184 North Ryde, Sydney N.S.W. Australia. (Australia) Tel. 61-2-886-4970

*To whom all correspondence and reprint requests should be directed

REFERENCES

1. Hershey, A.D. (Ed.) (1971) *The Bacteriophage Lambda*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (U.S.A.).
2. Hendrix, R.W., Roberts, J.W., Stahl, F.W., and Weisberg, R.A. (Eds.), (1983) *Lambda II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
3. Frischauf,A., Lehrach, H., Poustka, A., and Murray, N. (1983). *J. Mol. Biol.* 170: 827-842.
4. Karn, J., Brenner, S., and Barnett, L. (1983) *Meth. Enzymol.* 101:3-19.
5. Karn, J., Matthes, H. W., Gait,M.J., and Brenner, S. (1984) *Gene* 32:217-214.
6. Loenen, W. A. M., and Blattner, F. R. (1983). *Gene* 26:171-179.
7. Rimm, D. L., Horness, D., Kucera, J., and Blattner, F. R. (1980). *Gene* 12:301-309.
8. Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Furlong, L., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E., and Smithies, O. (1977) *Science* 196, 161-169.
9. Zimmerman, S. B., and Pheiffer, B. H. (1983). *PNAS* 80, 5852-5856.
10. Hayashi, K., Nakazawa, M., Ishizaki, Y., and Obayashi, A. (1985). *Nucl. Acids Res.* 13, 3261-3271.
11. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory.
12. Schroeder, J. L., and Blattner, F. R. (1982) *Nucl. Acids Res.*. Vol 10 pp. 69-84.
13. Blattner, F. R., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Richards, J. E., Slightom, J. L., Tucker, P. W., and Smithies, O. (1978). *Science* 202: 1279-1284.
14. Lis, J. T. (1980). *Meth. Enzymol.* 65: 347-353.
15. Leach, D. R., and Stahl, F. W. (1983). *Nature*, 305:448-451.
16. Wyman, A. R., Wolfe, L. B., and Botstein, D. (1985) *P.N.A.S.* 82:2880-2884.
17. Lobel, L. I., Patel, M., King, W., Nguyen-Huu, M. C., and Goff, S. P. (1985). *Science* 228:329-332.
18. Reik, W., Weiher, H., and Jaenisch, R. (1985). *P.N.A.S.* 82:1141-1145.
19. Nakamura, K., and Inouye, M. (1982). *EMBO J.* 1:771-775.
20. Vieira, J., and Messing, J. (1982). *Gene* 19:771-775
21. Rosenberg, M., and Shimitake, H. (1981). *Nature* 292:128-132.
22. Rao, R. N. (1984) *Gene* 32:247-250.
23. Karn, J., Brenner, S., Barnett, L., and Cesaereni, C. (1980) *P.N.A.S.* 77:5172
24. Leder, P., Tiemeier, D., and Enquist, L. (1977) *Science* 196:175.
25. Seed, B. (1983). *Nucl. Acids Res.* 11:2427-2447.
26. Saito, I., and Stark, G.R. (1986). *PNAS* 83:8668.
27. Hernandez, V. J., Edlind, T. D., Young, R. F., and Ihler, G. M. (1985). *Gene*, 33:363-365.
28. Williams, B. G., Blattner, F. R., Jaskunas, S., and Nomura, M. (1977) *J. Biol. Chem.* 252:344-7354.
29. Sanger, F., Coulson, A. R., Hong, G.-F., Hill, D. F. and Petersen, G. B. (1982). *J. Mol. Biol.* 162:301-302.
30. Buechel, D. E., Gronenborn, B., and Mueller-Hill, B. (1980) *Nature* 283:541-545.
31. Roberts, R. J., Akusjarvi, G., Alestrom, P., Gelinas, R. E., Gingeras,

T. R., Sciaky, D., and Petterson, U. (1986) in Adenovirus DNA W. Doerfler (ed.) Martinus Nijhoff Publishing, Boston.
32. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). *Gene*, 33:103-119.
33. New England Biolabs catalog. (1986-87) New England Biolabs, Beverly, MA.